心理科学进展 2023, Vol. 31, No. 9, 1676–1687
Advances in Psychological Science

https://doi.org/10.3724/SP.J.1042.2023.01676

言语产生中音节频率效应的认知机制：跨语言视角*

潘家冰 张清芳

(中国人民大学心理学系，北京 100872)

摘 要 在口语词汇产生过程中，音节频率效应(syllable frequency effect)指的是高频音节相对于低频音节的加工优势。综述了言语产生过程中音节频率效应的表现形式和理论基础，从影响因素、发生阶段和神经机制等角度阐述了印欧语系和汉语中音节频率效应的跨语言差异。基于口语产生的理论模型和合适单元假说，结合不同语言的固有特性，分析了音节频率效应存在跨语言差异的原因，提出了一个有关汉语口语词汇产生中音节作用机制的模型，为音节在汉语和字母语言口语产生过程中的作用提供了新的视角。未来应进一步探究音节频率的两种测量方式对音节频率效应产生影响的具体机制，加强对汉语口语词汇产生中音节频率效应的考察，利用多种实验技术和范式深入探讨言语产生过程中音节加工的跨语言差异。

关键词 言语产生，音节频率效应，跨语言，心理音节表，音节邻近项

分类号 B842

1 引言

言语产生在人类沟通交流中起着重要的作用，是指将思想转化为言语，并利用发音器官进行表达的心理过程，包含三个认知加工阶段：一是概念准备，即讲话者明确自己想要用言语表达的概念和内容；二是言语组织，即对概念进行组织并转换为相应的语言形式；三是发音运动，即通过声带肌肉运动进行发音输出(Levelt et al., 1999)。根据言语产生的 WEAVER++ (Word-form Encoding by Activation and Verification++, Levelt et al., 1999)模型，言语组织即词汇通达(lexical access)阶段，包括了词汇选择(lexical selection)、词素音位编码(morphological encoding)和音韵编码(phonological encoding)阶段。讲话者在音韵编码阶段根据词素提取音段和节律信息，并通过音节化过程(syllabification)将音段以严格的顺序插入节律框架中；之后进行语音编码(phonetic encoding)，提取音节的发音运动程序并为发音输出做好准备。已有研究对于音节在口语词汇产生中的作用机制存在争议。

音节是受具体语言结构和语义制约的自然发音单元(张清芳, 2005)。在言语产生过程中，音节加工的认知机制和时间进程存在跨语言差异。已有研究在汉语口语词汇产生中观察到了稳定的音节启动效应(syllable priming effect)，即启动刺激和目标刺激具有音节重叠时命名潜伏期更短。研究者认为这一效应表明音节在词汇选择后首先被提取，是音韵编码阶段重要的功能单元(Cai et al., 2020; Chen et al., 2002, 2016; Feng et al., 2019; Zhang & Damian, 2019; 蒋宇宸 等, 2020; 张清芳, 王雪娇, 2020)。而对于英语、荷兰语等印欧语系语言，人们对音节的通达和加工主要发生在语音编码阶段，表现为音节频率效应(syllable frequency effect, Bürki et al., 2015; Cholin, 2008; Laganaro & Alario, 2006)，在图画命名或词汇朗读过程中，高音节频率词语的命名潜伏期更短。目前，尽管言语产生过程中的音节频率效应已在不同语言当中得到了大量验证(例如 Croot et al., 2017; Laganaro & Alario, 2006; Levelt & Wheeldon, 1994; 杨群, 张清芳, 2015)，但缺乏对该领域研究成果的系统论述和跨语言比较。本文综述和比较印欧语系和汉语中音节频率效应的认知加工机制，以深入理

收稿日期：2022-12-22
* 教育部人文社科规划基金项目(21YJA190011)、国家语委"十四五"科研规划 2021 年度重点项目(ZDI145-6)和国家自然科学基金面上项目(32171055)资助。
通信作者：张清芳, E-mail: qingfang.zhang@ruc.edu.cn

解言语产生过程中音节频率效应的跨语言差异。下面我们首先介绍有关音节频率的测量方式，然后分别阐述印欧语系和汉语口语产生中音节频率效应的认知机制及相关理论。

2　音节频率的测量：类型频率和实例频率

音节频率的测量与音节邻近项(syllabic neighbours)密切相关。在字母语言中，音节邻近项广义上是指无论音节在哪个位置，相同位置上共享同一个音节的所有词(Perea & Carreiras, 1998)，狭义上是指仅在首音节上具有相同音节的所有词(Chetail & Mathey, 2011)，例如西班牙语词 casa (意思为“房子”，下同)、caro (“昂贵的”)、caja (“盒子”)和 cama (“床”)等互为音节邻近项(见 Perea & Carreiras, 1998)。汉语作为声调语言，存在大量的同音字和近音字(端木三, 2021; 史有为, 2019; 朱晓农, 2019)，汉语中的音节邻近项是指音节相同的所有单字(不考虑声调)。音节频率通常存在两种测量方式：音节邻近项个数和音节邻近项累计词频，前者被称为类型频率(type frequency)，即目标音节所拥有的全部音节邻近项的数量；后者被称为实例频率(token frequency)，是将该音节的所有音节邻近项的词频进行算数求和(Conrad et al., 2008)。例如，汉语音节/shua/的音节邻近项包括“刷”、“唰”和“耍”，根据汉语词频库 SUBTLEX-CH (Cai & Brysbaert, 2010)提供的词频，三者的词频分别为 16.4 次/百万、0.21 次/百万和 54.37 次/百万，因此，音节/shua/的类型频率为 3，实例频率为 70.98 次/百万。

已有多数研究仅考虑了音节的实例频率(例如 Bürki et al., 2015; Carreiras & Perea, 2004; Conrad et al., 2006; Croot et al., 2017; Farrell & Abrams, 2014; Hutzler et al., 2005; Laganaro & Alario, 2006; Perea & Carreiras, 1998; Stenneken et al., 2007)，同时测量了两种音节频率的研究很少，且本质上仍未真正区分二者所产生的效应。研究者同时测量了两种音节频率，但目的仅在于选取两种频率值具有可比性的音节作为实验材料，以确保高频音节同时具有较高的实例频率和类型频率，低频音节同时具有较低的实例频率和类型频率，而没有探讨这两种音节频率的效应是否有所不同(Cholin et al., 2006, 2011; Cholin & Levelt, 2009)。单词的实例频率和类型频率具有紧密的联系，研究中很难对这两个变量进行独立的操纵(Conrad et al., 2008; Miguel-Abella et al., 2022)，难以对音节频率效应进行明确归因。Conrad 等人(2008)认为这两种音节频率在不同的加工水平上起作用：音节邻近项累计词频容易受到词频极高的音节邻近项的影响，反映词汇水平上各音节邻近项和目标词之间的竞争，倾向于在词汇选择过程中产生抑制效应；相比之下，音节邻近项个数能够更好地反映亚词汇水平上音节表征本身的使用频率和典型性。Kwon (2014)采用韩语词汇朗读任务来检验上述观点。作为一种表音文字，韩语与印欧语系的语言类似，其正字法深度较浅，词汇朗读过程更多依赖拼读规则直接通达音节表征，而较少依赖词汇水平的信息，因此预期对音节表征敏感的类型频率会表现出促进效应，而对词汇水平敏感的实例频率不存在任何效应。结果发现类型频率和实例频率都表现出了促进效应，因此研究者认为二者均可以反映讲话者在亚词汇水平上对音节表征本身的加工。两种测量方式可能产生不同方向的效应，这提示研究者两类测量方式影响口语词汇产生过程的机制是不同的。

3　印欧语系中音节频率效应的认知机制

采用图画命名、词汇朗读或符号−词汇联想命名任务，研究者发现与命名低音节频率词相比，高音节频率词的命名潜伏期更短，即出现了音节频率效应(Cholin, 2008; Laganaro & Alario, 2006; Levelt & Wheeldon, 1994)。言语产生过程中的音节频率效应已在英语(Cholin et al., 2011; Croot et al., 2017; Macizo & Van Petten, 2007)、法语(Laganaro & Alario, 2006; Perret et al., 2014)、荷兰语(Cholin et al., 2006; Cholin & Levelt, 2009; den Hollander et al., 2019; Levelt & Wheeldon, 1994)、西班牙语(Carreiras & Perea, 2004; Onochie-Quintanilla et al., 2019; Perea & Carreiras, 1998)和德语(Conrad et al., 2006; Hutzler et al., 2005)等印欧语系语言中得到了重复验证。

3.1　音节频率效应的表现形式和发生阶段

目前，考察印欧语系言语产生过程的大多数研究发现了音节实例频率的促进效应(例如 Bürki et al., 2015; Carreiras & Perea, 2004; Cholin et al., 2011; Conrad et al., 2006; Croot et al., 2017; Laganaro & Alario, 2006; Perea & Carreiras, 1998)，但也有

少数研究报告了实例频率的抑制效应(Carreiras et al., 1993; Conrad et al., 2006; Farrell & Abrams, 2014; Hutzler et al., 2005)或无任何效应(Hutzler et al., 2005; Stenneken et al., 2007)。音节频率对言语产生过程不同方向的影响可能源于其对言语产生过程的两个加工阶段产生了不同的影响(Conrad et al., 2006; Hutzler et al., 2005; Perea & Carreiras, 1998)：在词汇通达中的词汇选择阶段，高音节频率词由于拥有更多的音节邻近项，能够激活更大的初始候选词群，且该词群更有可能包含高频词，这些激活的词汇节点之间相互竞争并干扰目标词的通达，故表现为抑制效应；而在语音编码阶段，讲话者从心理音节表(mental syllabary)中提取音节的发音运动程序。心理音节表是一个音节存储库，包括抽象的音韵音节(phonological syllable)和语境相关的语音音节(phonetic syllable)，后者即音节的发音运动程序。在语音编码阶段，音韵音节作为心理音节表的输入信号，映射到其中存储的发音运动程序，激活的发音运动程序将会从心理音节表中输出，进一步传输到发音系统以完成发音(Bürki et al., 2015; Cholin, 2008; Levelt, 1993; Levelt & Wheeldon, 1994)。由于高频音节的发音运动程序能够更快地从心理音节表中提取出来，使讲话者在构建高音节频率词的语音输出时更加容易，故表现为促进效应(见图 1a)。研究者通过让被试出声重复听到的单词，发现音位配列概率(probabilistic phonotactics，即音段和音段序列的频率)促进了亚词汇水平的加工，而相似性邻域密度(similarity-neighborhood density，即音韵邻近项累计词频)产生了抑制效应，这可能是由词汇水平上音韵邻近项(即由一个音素替换、添加或删除得到的词)之间竞争导致的(Vitevitch & Luce, 1998; 1999)。因此，在言语产生过程中与音韵信息相关的效应中，促进效应可能反映了词汇和亚词汇水平上由音韵表征本身特性导致的加工优势，而抑制效应则源于词汇水平上激活词条之间的竞争。

研究者在词汇判断任务中通常观察到音节频率的抑制效应(实例频率：Carreiras et al., 1993; Conrad et al., 2006, 2007, 2008; Conrad & Jacobs, 2004; 类型频率：Álvarez et al., 2001)，在假词命名任务中则通常观察到音节频率的促进效应(实例频率：Carreiras & Perea, 2004; Cholin et al., 2006, 2011; Conrad et al., 2006; Croot et al., 2017)，为口语词汇产生中音节频率的抑制效应和促进效应分别发生在词汇选择和词汇选择后的加工阶段(包括音韵编码和语音编码)提供了支持证据。词汇

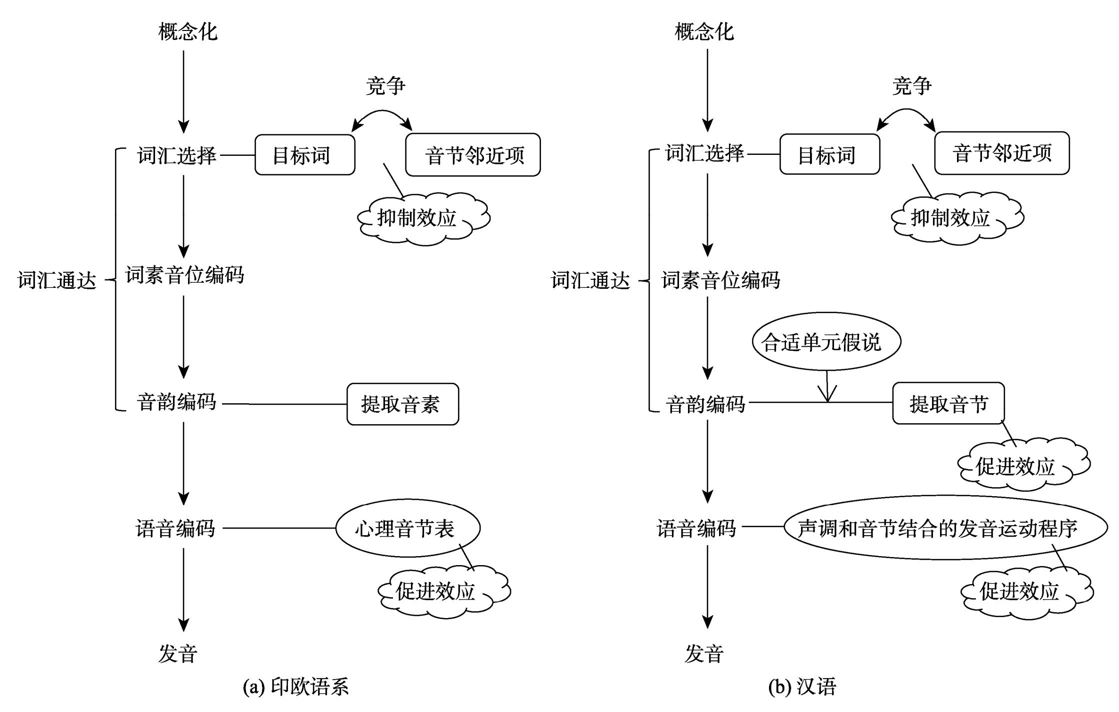


图 1 (a)印欧语系和(b)汉语口语词汇产生过程中的音节频率效应模型示意图

判断任务要求人们判断刺激是真词还是假词，所涉及的加工过程并不包含语音输出阶段，音节频率主要在词汇选择阶段起抑制作用。与低频音节相比，高音节频率的音节邻近项会产生较强的激活，与目标词之间形成竞争，干扰了目标词的真假词判断，出现了音节频率的抑制效应；而假词命名任务涉及的加工过程则不包含词汇选择阶段，因此音节频率主要在音韵编码、语音编码阶段起作用，高频音节相对于低频音节有提取优势，产生了促进作用。Hutzler 等人(2005)在真词命名任务中并未观察到音节频率效应，他们认为没有发现音节频率效应可能是由于音节频率对词汇选择的抑制作用和对语音输出的促进作用具有相等大小的强度，在反应时的结果上产生了相互抵消。在言语产生过程中，高低频音节对词汇选择的抑制作用和对语音输出的促进作用之间的相互权衡共同决定了音节频率效应的具体表现形式。

根据口语词汇产生过程的 WEAVER++模型，在音韵编码阶段讲话者首先提取音段和节律信息，然后通过音节化过程构建音节结构，并在语音编码阶段进一步提取音节的发音运动程序(Levelt et al., 1999)。对于音节频率所表现出来的促进效应，已有研究对其发生在音韵编码还是语音编码阶段存在争论。研究者主要采用了两种研究思路：一种研究思路是基于相加因素法(additive-factor method, Sternberg, 1969)，检验音节频率与其他语言因素的效应之间是否存在交互作用。Cholin 和 Levelt (2009)采用内隐启动范式观察到荷兰语音节准备效应和音节频率效应之间存在交互作用。由于音节准备效应包含了单词形式编码的所有加工阶段(即词素音位编码、音韵编码和语音编码)，根据相加因素法的原理，两个因素之间存在交互作用表明二者可能作用于序列加工过程中相同的阶段，因此研究者认为该结果表明心理音节表可能位于整词通达之后的语音编码阶段。第二种研究思路是采用包含不同口语词汇产生阶段的任务，通过比较包含不同阶段的任务间音节频率效应是否出现推断其发生阶段。Laganaro 和 Alario (2006)采用即时命名任务、延时命名任务以及有发音抑制的延时命名任务来考察法语口语词汇产生中音节频率效应的发生阶段。在三种任务中，即时命名任务包含从概念化到发音的全部言语产生过程；延时命名任务并不要求被试在看到图片后立刻命名，而是只有当屏幕上出现提示符时才开始出声命名，在发音前被试已经完成了对目标词的音韵编码和语音编码，当提示线索出现时执行已经准备好的发音运动程序即可，因此该任务的命名潜伏期仅反映了发音阶段的加工；有发音抑制的延时命名任务则要求被试在图片出现后一直出声重复某个特定语音(例如/da/)，当看到提示符时立即停止并完成图片命名。由于重复发音将占用讲话者的发音环路(articulatory loop, Baddeley et al., 1984)并阻断对目标词的语音编码，因此提示符出现后所测量的命名潜伏期反映了语音编码和发音两个阶段的加工。研究者在即时命名任务以及有发音抑制的延时命名任务中均观察到了音节频率效应，而在延时命名任务中未观察到。即时命名和有发音抑制的延时命名任务中均包括语音编码过程，而延时命名任务中未包括，任务之间的对比结果表明音节频率效应发生在语音编码阶段。然而，Croot 等人(2017)采用类似的实验设计来考察英语口语词汇产生时，仅在即时命名任务中发现了音节频率效应。根据上述实验任务间的对比，该结果表明音节频率效应并不发生在语音编码和发音阶段，而是可能发生在音韵编码阶段。Croot 等(2017)指出音节频率效应也有可能出现在语音编码阶段，只是效应比较微弱而未能在有发音抑制的延时命名任务中检测到。

高时间分辨率的技术能够精细地考察到音节频率效应发生的时间进程。Bürki 等人(2015)利用事件相关电位技术发现，法语高频音节和新异音节(音节频率为 0)之间诱发的波幅差异出现在发音前的 180~160 ms 以及 115~100 ms。口语词汇产生过程中与该时间窗相对应的加工阶段是语音编码(Indefrey, 2011; Laganaro et al., 2013)，脑电研究的发现表明音节频率效应发生在语音编码阶段。综上所述，多数研究证据支持印欧语系中的音节频率效应发生在语音编码阶段，但目前也并不能排除音节频率效应发生在音韵编码阶段的可能性。音节频率的促进效应可能出现在音韵编码或(和)语音编码阶段，尽管英语和法语同属于印欧语系，然而其语言系统中词汇和音韵之间的关系也存在差异，跨语言之间的差异可能会影响口语词汇产生中音节频率效应的认知机制。

除此之外，研究者也关注了音节的位置特征对印欧语系言语产生中音节频率效应的影响。对

于双音节词和多音节词，音节的位置频率(positional frequency)是指音节在单词中的特定位置上所出现的次数(以词频加权，Carreiras & Perea, 2004; Perea & Carreiras, 1998)，或同一位置上包含该音节的所有单词的累计词频(Farrell & Abrams, 2014; Hutzler et al., 2005)。研究发现排除双音节词重音分配模式的潜在影响后，音节的位置会影响音节频率效应的有无，但已有研究结果仍存在争议。Levelt 和 Wheeldon (1994)在荷兰语双音节真词的命名中仅观察到了第二音节的频率效应，研究者认为只有当单词中全部音节都完成语音编码后才会输出发音，因此先提取出的高频首音节在加工速度上的优势变得不明显。然而也有研究在西班牙语(Carreiras & Perea, 2004)和荷兰语(Cholin et al., 2006)的双音节假词命名中仅观察到了首音节频率效应，表明当首音节的语音编码完成后就可以立刻开始发音，即对首音节进行发音输出的同时检索提取第二音节，因此在快速的命名任务中第二音节对频率操纵不敏感。Cholin 等人(2011)提出上述研究之间讲话者采取不同语音检索策略的原因可能在于实验材料的词汇性。在真词口语产生中，人们会在整个单词完成语音编码后才开始发音以正确地传达单词意义，而对于假词没有传达意义的要求，人们只需要发音即可。也有针对英语双音节词命名的研究同时发现了首音节和第二音节的频率效应(假词命名：Cholin et al., 2011; 真词命名：Macizo & Van Petten, 2007)，反映了英语母语者在语音编码和发音的界面上可能具有更大的计划广度。Cholin 等(2011)提出音节边界的清晰性会影响不同语言的发音计划广度：与英语相比，荷兰语和西班牙语的音节边界相对清晰，讲话者能够较早地检测到音节边界，音节化过程以及通达心理音节表的速度也较快。在快速命名任务中，荷兰语和西班牙语的讲话者倾向于采用音节大小的发音计划广度，而英语中音节边界相对模糊，检测到音节边界较为困难，讲话者倾向于采用跨音节边界的、多音节的发音计划广度。多音节词汇产生中不同位置上的音节是否表现出频率效应，不仅与讲话者在命名真词和假词时依赖不同的语音检索策略有关，而且与发音计划广度的跨语言差异有关。

3.2 音节频率促进效应的认知机制：理论观点

研究者对于口语产生中音节频率促进效应的发生机制存在争论，提出了心理音节表理论(mental syllabary theory, Levelt, 1993)、混合模型(mixed model, Levelt & Wheeldon, 1994)和双通路模型(dual-route model, Varley & Whiteside, 2001; Whiteside & Varley, 1998)。

Levelt (1993)提出了心理音节表理论(mental syllabary theory)来解释字母语言口语词汇产生中的音节频率促进效应。根据该理论，音节频率效应源于讲话者直接提取心理音节表中存储的发音运动程序，由于高频音节比低频音节能够更快地通达心理音节表，因此个体对高音节频率词语的命名更快(Cholin et al., 2006; Cholin & Levelt, 2009; Ferrand et al., 1996; Levelt et al., 1999)。心理音节表理论的意义在于提取音节的高效性，由于大多数音节是多次使用的发音单元，若每次使用时都重新计算其完整的发音运动程序会非常浪费加工资源(Levelt, 1993)，而通过直接提取存储在心理音节表中的发音运动程序，将大大减少言语产生系统的计算负荷并有助于快速流畅地发音(Carreiras & Perea, 2004; Cholin, 2008; Hagoort et al., 1999)。

Levelt 和 Wheeldon (1994)通过荷兰语口语词汇产生中的音节频率效应验证了心理音节表理论，同时他们指出讲话者可以正确读出由不存在的音节组成但音韵结构符合规则的假词(例如 fliltirp)。心理音节表中实际并未存储这种不存在的音节(即新异音节)，因此讲话者无法通达心理音节表来提取其发音运动程序，对新异音节的加工可能依赖于其他机制。因此他们进一步提出了混合模型(mixed model)，假设在心理音节表中存在直接提取音节和在线计算音节两种加工方式：新异音节和极低频音节采取在线计算的方式，常用的高频音节则存储在心理音节表中直接提取(同见 Bürki, 2018)。音节频率促进效应源于直接提取高频音节的速度快于在线计算低频音节的速度。Ferrand 等人(1996)以法语中的音节数量为例，指出并非所有音节都存储在心理音节表中，语言中潜在的新异音节可能是通过在线即时组装而来的：法语中实际存在大约 6000 个音节，若同时纳入其他不符合法语单词形式规则但仍然遵循法语发音规则的潜在音节(例如/kmƐlu/)则可能多达 33600 个，讲话者没有必要在心理音节表中存储如此大量的音节，可以通过类比的方式对符合音

韵规则的新异音节产生正确发音。正如在语音习得过程的早期, 所有音节都相当于新异音节, 讲话者是在习得过程中依据语言使用经验才逐步形成从心理音节表中提取音节的加工机制。

语音编码的双通路模型(dual-route model, Varley & Whiteside, 2001; Whiteside & Varley, 1998)同样认为高频音节和低频音节依赖于不同的加工机制：高频音节通过直接通路从心理音节表中进行提取, 低频音节通过间接通路在线组装音段形成(同见 Cholin, 2008; Cholin & Levelt, 2009; Hagoort et al., 1999; Tremblay et al., 2018)。来自首音段替换错误的证据支持了音段在线组装成音节的观点：在连续语流中两个单词的首音段被交换位置(例如 car park 被错误发音为/pa: ka:k/), 这是由于音段在分配给音节框架对应槽的过程中出现了错误, 表明音段可以作为言语组装过程中的独立表征单元(Varley & Whiteside, 2001)。双通路模型进一步假设音节的提取与其静息激活水平(resting activation level)有关, 高频音节的静息激活水平较高, 从心理音节表中进行选择和提取出来的阈值较低(Bürki et al., 2015), 因此直接提取高频音节要比在线组装低频音节更容易。作为对心理音节表理论的补充和发展, 混合模型和双通路模型同样预测讲话者对高音节频率词的命名快于低频音节, 表现为促进效应, 但促进效应出现的认知机制是不同的。

3.3　音节频率效应的神经机制

近年来, 神经影像学研究利用多种成像技术考察了印欧语系中音节频率效应的神经机制。基于混合模型和双通路模型的观点, 高频音节和低频音节的发音运动程序与不同的认知加工机制相关联, 因此在产生不同频率的音节时, 大脑的激活模式可能不同。Hagoort 等人(1999)利用 PET 技术考察了德语母语者对真词(包含更多高频音节)和假词(包含更多低频音节)的默读和出声朗读过程, 由于该研究混淆了刺激的音节频率和词汇性, 因此研究者仅推测辅助运动区可能参与了高频音节通达心理音节表的过程, 而左内侧前运动皮层可能参与了低频音节的在线组装过程。

已有利用 fMRI 技术的研究发现, 低频音节相比于高频音节能够引起更加强烈的脑区激活, 而相反方向的对比中则不存在显著差异, 表明产生低频音节需要更高的加工成本(Bürki et al., 2015)。例如, Carreiras 等人(2006)在西班牙语词汇朗读任务中发现, 低音节频率词相比于高音节频率词显著激活了左侧前脑岛, 而高音节频率词与低音节频率词相比则没有引起特定脑区的显著激活, 研究者指出低频音节引起的激活更强可能与其通达存储的运动程序的速度更慢有关, 而高频音节则能够促进发音运动的计划过程, 支持了心理音节表理论的观点。类似地, Papoutsi 等人(2009)要求英语母语者听音节频率不同的假词并在 6 秒的延迟后重复它们, 结果发现低音节频率词显著激活了前运动皮层中的背侧中央前回、双侧额下回和辅助运动区, 而高音节频率条件与低音节频率条件相比同样没有观察到显著激活的脑区。然而, 上述研究所发现的高频音节和低频音节在脑区激活程度上的差异无法证明二者的加工是由不同脑区控制的, 虽然高频音节与低频音节相比没有诱发特定脑区的激活, 但二者仍可能存在神经网络上的差异, 只是差异发生的时间窗口太窄, 无法用 fMRI 等低时间分辨率的成像技术检测到。进一步地, Bürki 等人(2015)采用 ERP 技术考察了包含高频、低频和新异音节的法语假词命名, 结果显示在发音开始前的 170~100 ms 左右, 高频和低频/新异音节之间的 ERPs 波幅和整体地形模式上出现了差异, 而后两类音节具有相同且稳定的电生理模式, 表明个体在产生高频音节和低频/新异音节时依赖不同的神经机制, 而在产生低频音节和新异音节时依赖相似的神经机制。研究者认为该结果反映了高频音节的加工优势可能是源于其发音运动程序存储在心理音节表中, 而低频音节的加工机制则类似于新异音节, 其发音运动程序是通过在线组装的方式生成的, 为音节频率的双通路模型提供了支持证据(Bürki et al., 2015)。

4　汉语中音节频率效应的认知机制

4.1　音节在汉语口语词汇产生中的作用

作为一种声调语言和非拼音文字, 汉语和字母语言的语言特点存在很大差异。字母语言不仅音节数量众多(例如荷兰语和英语中均存在超过 12000 个音节), 而且存在模糊音节和重新音节化(resyllabification)现象, 即连续语流中前一词的词尾辅音可以跨越词汇边界与后一词的词首元音重新组合成音节(Levelt et al., 1999)。与字母语言不同, 汉语中音节数量很少, 不考虑声调大约仅有

400 个，即使考虑声调也仅有 1200 个左右，且汉语中音节之间界限清晰，不存在重新音节化的现象(张清芳，杨玉芳，2005)。因此，在音韵编码阶段，汉语母语者可以直接从心理词典中提取整个音节(O' Seaghdha et al., 2010; 蒋宇宸 等，2020; 张清芳，王雪娇，2020)。

音节在口语词汇产生中的作用存在跨语言差异。对于印欧语系的语言，音韵编码阶段首先提取的加工单元是音素(Alario et al., 2007; Damian & Bowers, 2003; Damian & Dumay, 2007; Jacobs & Dell, 2014; Meyer, 1991; Schiller, 2008)。音素(phoneme)是能够区分意义的最小语音单元，分为元音和辅音(Clark & Yallop, 1995)。根据WEAVER++模型，讲话者随后通过音节化过程来构建音节(Levelt et al., 1999)，完成语音编码。在汉语中，采用内隐启动范式(Chen et al., 2002)、掩蔽启动范式(Chen et al., 2016; You et al., 2012)和图词干扰范式(岳源，张清芳，2015；张清芳，杨玉芳，2005)等，大量研究发现音节是汉语口语产生过程中的音韵编码单元。有研究利用 ERP 技术比较音节效应和音素效应的时间进程，结果发现汉语口语产生中音节的提取早于音素，音节在音韵编码阶段发挥作用，而音素效应主要出现在音韵编码阶段的晚期或语音编码阶段(Cai et al., 2020; Feng et al., 2019; Zhang & Damian, 2019; 张清芳，王雪娇，2020)；但也有研究在音韵编码和语音编码两个阶段均观察到了音节效应(Wang et al., 2017)。特别地，蒋宇宸等人(2020)通过时频分析发现在音韵编码阶段出现了音节效应而非音素效应的 θ 频段能量变化，从神经振荡的角度证实了音节是汉语口语产生过程中的音韵编码单元。近年来，尽管有研究发现音素在汉语口语产生的音韵编码阶段也会激活(Qu et al., 2012, 2020; Yu et al., 2014)，但目前有关汉语音素加工方面的证据相对较少且存在争议(屈青青 等，2018)，已有研究主要证实了音节在汉语音韵编码中的重要作用。

基于印欧语系和汉语中的不同研究结果，研究者提出了合适单元假说(proximate units principle, O' Seaghdha et al., 2010)来阐述言语产生过程中音韵编码单元的跨语言差异。合适编码单元(proximate units)是指词素信息被激活之后最先选择的音韵加工单元，印欧语系中的合适编码单元是音素，汉语中则是音节。对于印欧语系的语言，首先提取完音素后需要结合节律框架进行音节化过程，随后在语音编码阶段从心理音节表中提取音节的发音运动程序。与此不同的是，汉语的音韵编码过程中讲话者首先提取音节，再进一步将音节分解为音段信息(即音素或大于音素而小于音节的音韵组合单元)和超音段信息(即声调)，然后完成语音编码和发音过程(见图 2)。Roelofs (2015)通过计算模拟的方法对该假说进行了验证。

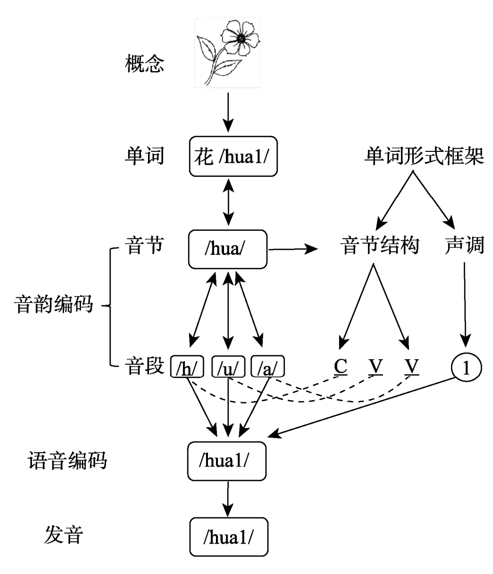


图 2 合适单元假说中汉语单音节词汇的口语产生过程

4.2 音节频率效应的表现形式和发生阶段

已有汉语音节频率效应的研究主要采用了图画口语命名任务，要求讲话者用单字词(即单音节词)对图画进行命名(例如 Zhang & Wang, 2014; 杨群，张清芳，2015)。大多数研究均观察到了实例频率的促进效应(例如 Zhang & Wang, 2014; 欧阳明昆，2020; 杨群，张清芳，2015)，而忽略了对类型频率的考察。仅有一项研究(蒋宇宸，2021)采用图画命名和单字命名任务发现，当同时操纵实例频率和类型频率时，在图画命名和单字命名中分别发现了音节频率的促进效应和抑制效应；当匹配类型频率操纵实例频率时，图画命名中的音节频率促进效应仍然存在，但单字命名中未发现任何频率效应。作为典型的口语词汇产生任务，图画口语命名中发现的音节频率促进效应，效应方向与印欧语系的研究一致，而且研究进一步发现主要是由于实例频率的变化引起了音节频率促进效应。而单字命名任务中不同变量操纵下的效应

对比表明音节的类型频率是引起音节频率效应的主要因素。单字命名中音节频率的抑制效应可能源于高音节频率词会激活更多的音节邻近项，以并行分布加工(parallel distributed processing, Seidenberg & McClelland, 1989)的方式相互作用并对目标词造成干扰，因此高音节频率词的命名潜伏期更长(蒋宇宸, 2021)。此外，在单字命名任务中，视觉呈现目标单词要求被试大声读出单词的任务中既包括了视觉再认单词过程，同时包括了口语产生中的音韵编码、语音编码和发音过程。单字命名过程中所发现的抑制效应可能是由于视觉词汇通达阶段所产生的抑制效应与音韵编码和语音编码过程中所产生的促进效应互相抵消后的结果。这一发现表明音节的类型频率和实例频率在汉语口语词汇产生中起不同作用，然而目前的研究尚不能清晰地阐述两者的具体作用。

如前所述，基于合适单元假说和有关汉语口语产生中音节加工的研究成果，汉语母语者在音韵编码过程中提取词汇的音节信息(Cai et al., 2020; Feng et al., 2019; O' Seaghdha et al., 2010; Zhang & Damian, 2019)，因此汉语中的音节频率效应很有可能发生在音韵编码阶段。在行为层面上，Zhang 和 Wang (2014)采用汉语图画命名任务并未观察到词频和音节频率之间的交互作用，表明这两个因素独立影响了命名过程，与印欧语系中的结果(Levelt & Wheeldon, 1994)相一致。而有研究采用图词干扰范式观察到了音节频率和音韵相关性之间的交互作用(欧阳明昆, 2020; 杨群, 张清芳, 2015)，汉语口语产生过程中由音韵相关性引起的音节促进效应发生在音韵编码阶段(岳源, 张清芳, 2015)，因此该结果表明汉语中的音节频率效应很有可能也发生在音韵编码阶段。然而，有研究采取了相同的图词干扰范式却未能成功重复出上述结果，研究者认为该结果表明汉语口语词汇产生中的音节频率效应并不发生在音韵编码阶段，而是可能发生在语音编码阶段(蒋宇宸, 2021)。在时间进程方面，杨群(2017)借助 ERP 技术发现青年人的音节频率效应主要出现在 250~350 ms 之间，根据口语词汇产生时间进程的元分析结果，该时间窗对应于音韵编码过程(Indefrey, 2011; Indefrey & Levelt, 2004)，因此该结果进一步证实了汉语音节频率效应发生在音韵编码阶段。综上所述，尽管目前尚不能排除音节频率也有可能对汉语的语音编码阶段产生影响，但已有研究成果大多支持汉语音节频率效应发生在音韵编码阶段(见图 1b)，为音节在汉语口语产生的音韵编码阶段发挥重要作用提供了支持证据。

根据已有研究发现，我们提出了一个有关汉语口语词汇产生中音节作用机制的模型(图 1b)，为后续研究提供理论基础。在词汇选择阶段，目标词与各音节邻近项形成相互竞争，与低频音节相比，高频音节的邻近项激活程度更高，产生词汇水平上的抑制效应。在音韵编码阶段，讲话者提取音节进行音韵编码过程，高频音节的提取快于低频音节，产生音节频率促进效应。语音编码阶段讲话者针对所选定的目标词进行发音运动程序上的编码，在此过程中必须考虑到音节和声调的结合。高频音节对应的发音运动程序由于多次使用其提取也会快于低频音节，因而在此阶段也有可能产生音节频率的促进效应。最终完成发音输出时表现为音节频率促进效应还是抑制效应，依赖于各个加工阶段所产生的抑制效应和促进效应的大小，以及各个加工阶段之间的关系(独立的还是交互的)。目前有关汉语音节频率效应的研究相对较少，未来还需要借助多个实验任务和技术，在行为、电生理和脑机制层面深入考察汉语音节频率效应的特征和认知机制并进行跨语言比较。

5　跨语言视角下的音节：总结与展望

音节频率效应广泛存在于人类语言认知加工中，研究者通过图片命名、词汇朗读、假词命名和图词干扰等实验范式，从跨语言的视角系统探讨了印欧语系和汉语的言语产生过程中音节频率效应的认知加工机制。音节频率效应的跨语言差异主要体现在表现形式和发生阶段两个方面。在表现形式上，印欧语系语言中发现了音节频率促进效应、抑制效应或无任何效应，而汉语中则是以音节频率的促进效应为主。音节频率效应的方向与所操纵的频率指标(类型频率或实例频率)和任务密切相关。在发生阶段上，尽管目前无论是印欧语系还是汉语的研究对于音节频率效应的发生阶段均存在争议，但已有理论和研究证据多数支持印欧语系口语词汇产生中的音节频率效应发生在语音编码阶段，而汉语口语词汇产生中的音节频率效应可能发生在音韵编码和语音编码阶段。音节频率效应的跨语言差异可能与不同语言

系统的固有特性有关，为音节在字母语言和汉语口语产生过程的不同加工阶段发挥作用提供了新的视角。根据合适单元假说，汉语的音韵编码单元是音节(Cai et al., 2020; Chen et al., 2002, 2016; Feng et al., 2019; Zhang & Damian, 2019; 蒋宇宸等, 2020; 张清芳, 王雪娇, 2020)，汉语中音节数量较少，因此在音韵编码阶段讲话者可以直接提取音节，高频音节的提取快于低频音节，产生音节频率促进效应，音节频率和图词干扰范式中的语音相关性之间的交互作用为音节频率效应发生在音韵编码阶段提供了支持证据(欧阳明昆, 2020; 杨群, 2017; 杨群, 张清芳, 2015)。根据语音编码阶段的特点，高频音节在音韵编码阶段提取较快，其发音运动程序的编码和提取也会快于低频音节，我们认为音节频率效应也可能发生在语音编码阶段。基于上述研究现状，汉语口语词汇产生中的音节频率效应亟需系统深入的研究。

具体地可以从以下方面展开深入探索：第一，有关类型频率和实例频率的作用需要进一步研究，两类测量指标是否在不同词汇产生任务中起了不同的作用，其作用可能也存在跨语言的差异。第二，音节频率促进效应的认知机制有待进一步探讨。目前关于音节频率效应的主流观点认为音节的加工是通过心理音节表起作用的(Laganaro, 2019)，混合模型和双通路模型进一步发展了心理音节表理论，对于高频音节和低频音节的加工机制持分离的观点，但均预测高音节频率词的命名潜伏期要短于低音节频率词，因此仅根据以往研究所得到的音节频率效应无法确定低频音节是存储在心理音节表中还是在线组装而成。目前除了Bürki等人(2015)通过高频音节和低频音节的神经电生理模式存在分离而支持了双通路模型的观点，其余还未有研究能够直接检验混合模型和双通路模型。由于汉语的语音系统特性不同于印欧语系，其音节频率效应的机制与印欧语系语言可能完全不同，需要更为深入系统的探索。第三，目前有关汉语音节频率效应的研究相对较少，已有研究关注的均是汉语单字词(即单音节词)的音节频率效应，而忽略了对双音节词或多音节词的考察，因此目前缺乏音节位置影响汉语音节频率效应的证据；此外，考察汉语音节频率效应发生阶段的研究所采用的实验任务较为单一(均为图词干扰范式)，未来还需着重关注不同的言语产生任务所涉及的认知加工过程，结合更加多元的言语产生实验范式，进一步明确汉语音节频率效应的发生阶段和神经机制，为探讨音节在汉语言语产生中的作用以及音节频率效应认知神经机制的跨语言差异提供更深入的理解。第四，从音节频率的视角思考汉语语言习得。由于汉语中存在大量的同音字和近音字，尤其是对于类型频率较高的音节来说，其拥有的大量音节邻近项使得人们在听到单一音节时通常难以通达一个确定汉字的语义，除非依赖语境信息的支持，因此汉语学习者往往需要在音节水平上整体记忆汉字的音韵形式，而非像学习字母语言的音韵形式那样可以依赖拼读规则。这启示汉语初学者可以通过书写或过度学习等策略来帮助区分同音字等易混淆的音节邻近项，从而在词汇选择过程中逐步降低易混淆音节邻近项的干扰程度，实现高音节频率词在语音构建上的优势。

参考文献

端木三. (2021). 英汉音节分析及数量对比. *语言科学, 20*(6), 561−588.

蒋宇宸. (2021). *音节累计词频和邻近项个数影响汉语口语词汇产生过程的认知机制* (硕士学位论文). 中国人民大学, 北京.

蒋宇宸, 蔡笑, 张清芳. (2020). θ 频段(4~8 Hz)的活动反映了汉语口语产生中音节信息的加工. *心理学报, 52*(10), 1199−1211.

欧阳明昆. (2020). *汉语词汇产生过程中语义和语音提取年老化的认知神经机制* (博士学位论文). 中国人民大学, 北京.

屈青青, 刘维琳, 李兴珊. (2018). 汉语言语产生的语音加工单元——基于音位的研究. *心理科学进展, 26*(9), 1535−1544.

史有为. (2019). 带调音节: 解释汉语的一个出发点. *华东师范大学学报 (哲学社会科学版), 51*(3), 1−17.

杨群. (2017). *汉语口语词汇产生中词汇通达过程的认知年老化机制* (博士学位论文). 中国人民大学, 北京.

杨群, 张清芳. (2015). 口语产生中词频效应、音节频率效应和语音促进效应的认知年老化. *心理科学, 38*(6), 1303−1310.

岳源, 张清芳. (2015). 汉语口语产生中音节和音段的促进和抑制效应. *心理学报, 47*(3), 319−328.

张清芳. (2005). 音节在语言产生中的作用. *心理科学进展, 13*(6), 752−759.

张清芳, 王雪娇. (2020). 汉语口语词汇产生的音韵编码单元: 内隐启动范式的 ERP 研究. *心理学报, 52*(4), 414−425.

张清芳，杨玉芳. (2005). 汉语单音节词汇产生中音韵编码的单元. *心理科学, 28*(2), 374−378.

朱晓农. (2019). 声调发生的五项前提. *语言科学, 18*(6), 561−580.

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition, 102*(3), 464−475.

Álvarez, C. J., Carreiras, M., & Taft, M. (2001). Syllables and morphemes: Contrasting frequency effects in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 545−555.

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A, 36*(2), 233−252.

Bürki, A. (2018). Variation in the speech signal as a window into the cognitive architecture of language production. *Psychonomic Bulletin & Review, 25*, 1973−2004.

Bürki, A., Cheneval, P. P., & Laganaro, M. (2015). Do speakers have access to a mental syllabary? ERP comparison of high frequency and novel syllable production. *Brain and Language, 150*, 90−102.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE, 5*(6), Article e10729. https://doi.org/10.1371/journal.pone.0010729

Cai, X., Yin, Y. L., & Zhang, Q. F. (2020). The roles of syllables and phonemes during phonological encoding in Chinese spoken word production: A topographic ERP study. *Neuropsychologia, 140*, Article 107382. https://doi.org/10.1016/j.neuropsychologia.2020.107382

Carreiras, M., Álvarez, C. J., & De Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32*(6), 766−780.

Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Human Brain Mapping, 27*(12), 963−972.

Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language, 90*(1−3), 393−400.

Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46*(4), 751−781.

Chen, J.-Y., O' Seaghdha, P. G., & Chen, T.-M. (2016). The primacy of abstract syllables in Chinese word production. *Journal of Experimental Psychology: Learning, Memory and Cognition, 42*(5), 825−836.

Chetail, F., & Mathey, S. (2011). Effect of syllabic neighbourhood as a function of syllabic neighbour length. *Psychonomic Bulletin & Review, 18*(5), 951−957.

Cholin, J. (2008). The mental syllabary in speech production: An integration of different approaches and domains. *Aphasiology, 22*(11), 1127−1141.

Cholin, J., Dell, G. S., & Levelt, W. J. (2011). Planning and articulation in incremental word production: Syllable-frequency effects in English. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 109−122.

Cholin, J., & Levelt, W. J. (2009). Effects of syllable preparation and syllable frequency in speech production: Further evidence for syllabic units at a post-lexical level. *Language and Cognitive Processes, 24*(5), 662−684.

Cholin, J., Levelt, W. J., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition, 99*(2), 205−235.

Clark, J., & Yallop, C. (1995). *An Introduction to Phonetics and Phonology* (2nd ed.). Oxford: Blackwell.

Conrad, M., Carreiras, M., & Jacobs, A. M. (2008). Contrasting effects of token and type syllable frequency in lexical decision. *Language and Cognitive Processes, 23*(2), 296−326.

Conrad, M., Grainger, J., & Jacobs, A. M. (2007). Phonology as the source of syllable frequency effects in visual word recognition: Evidence from French. *Memory & Cognition, 35*(5), 974−983.

Conrad, M., & Jacobs, A. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes, 19*(3), 369−390.

Conrad, M., Stenneken, P., & Jacobs, A. M. (2006). Associated or dissociated effects of syllable frequency in lexical decision and naming. *Psychonomic Bulletin & Review, 13*(2), 339−345.

Croot, K., Lalas, G., Biedermann, B., Rastle, K., Jones, K., & Cholin, J. (2017). Syllable frequency effects in immediate but not delayed syllable naming in English. *Language, Cognition and Neuroscience, 32*(9), 1119−1132.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language, 49*(1), 119−132.

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language, 57*(2), 195−209.

den Hollander, J., Jonkers, R., Mariën, P., & Bastiaanse, R. (2019). Identifying the speech production stages in early and late adulthood by using electroencephalography. *Frontiers in Human Neuroscience, 13*, Article 298. https://doi.org/10.3389/fnhum.2019.00298

Farrell, M. T., & Abrams, L. (2014). Picture−word interference reveals inhibitory effects of syllable frequency on lexical

selection. *Quarterly Journal of Experimental Psychology, 67*(3), 525−541.

Feng, C., Yue, Y., & Zhang, Q. F. (2019). Syllables are retrieved before segments in the spoken production of Mandarin Chinese: An ERP Study. *Scientific Reports, 9*(1), Article 11773. https://doi.org/10.1038/s41598-019-48033-3

Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllabic units. *Journal of Memory and Language, 35*(5), 708−723.

Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., & Seitz, R. J. (1999). The neural circuitry involved in the reading of German words and pseudowords: A PET study. *Journal of Cognitive Neuroscience, 11*(4), 383−398.

Hutzler, F., Conrad, M., & Jacobs, A. M. (2005). Effects of syllable-frequency in lexical decision and naming: An eye-movement study. *Brain and Language, 92*(2), 138−152.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology, 2*, Article 255. https://doi.org/10.3389/fpsyg.2011.00255

Indefrey, P. & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*, 101−144.

Jacobs, C. L., & Dell, G. S. (2014). 'Hotdog', not 'Hot' 'dog': The phonological planning of compound words. *Language, Cognition and Neuroscience, 29*(4), 512−523.

Kwon, Y. (2014). The syllable type and token frequency effect in naming task. *Korean Journal of Cognitive Science, 25*(2), 91−107.

Laganaro, M. (2019). Phonetic encoding in utterance production: A review of open issues from 1989 to 2018. *Language, Cognition and Neuroscience, 34*(9), 1193−1201.

Laganaro, M., & Alario, F. X. (2006). On the locus of the syllable frequency effect in speech production. *Journal of Memory and Language, 55*(2), 178−196.

Laganaro, M., Python, G., & Toepel, U. (2013). Dynamics of phonological−phonetic encoding in word production: Evidence from diverging ERPs between stroke patients and controls. *Brain and Language, 126*(2), 123−132.

Levelt, W. J. M. (1993). Timing in speech production with special reference to word form encoding. *Annals of the New York Academy of Sciences, 682*, 283−295.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1−38.

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition, 50*(1−3), 239−269.

Macizo, P., & Van Petten, C. (2007). Syllable frequency in lexical decision and naming of English words. *Reading and Writing, 20*(4), 295−331.

Meyer, A. S. (1991). The time course of phonological encoding in language production: Phonological encoding inside a syllable. *Journal of Memory and Language, 30*(1), 69−89.

Miguel-Abella, R. S., Pérez-Sánchez, M. Á., Cuetos, F., Marín, J., & González-Nosti, M. (2022). SpaVerb-WN—A megastudy of naming times for 4562 Spanish verbs: Effects of psycholinguistic and motor content variables. *Behavior Research Methods, 54*(6), 2640−2664.

Onochie-Quintanilla, E., Defior, S. A., & Simpson, I. C. (2019). RAN and orthographic processing: What can syllable frequency tell us about this relationship?. *Journal of Experimental Child Psychology, 182*, 1−17.

O' Seaghdha, P. G., Chen, J. Y., & Chen, T. M. (2010). Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition, 115*(2), 282−302.

Papoutsi, M., de Zwart, J. A., Jansma, J. M., Pickering, M. J., Bednar, J. A., & Horwitz, B. (2009). From phonemes to articulatory codes: An fMRI study of the role of Broca's area in speech production. *Cerebral Cortex, 19*(9), 2156−2165.

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 134−144.

Perret, C., Schneider, L., Dayer, G., & Laganaro, M. (2014). Convergences and divergences between neurolinguistic and psycholinguistic data in the study of phonological and phonetic encoding: A parallel investigation of syllable frequency effects in brain-damaged and healthy speakers. *Language, Cognition and Neuroscience, 29*(6), 714−727.

Qu, Q. Q., Damian, M. F., & Kazanina, N. (2012). Sound-sized segments are significant for Mandarin speakers. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 14265−14270.

Qu, Q. Q., Feng, C., Hou, F. Y., & Damian, M. F. (2020). Syllables and phonemes as planning units in Mandarin Chinese spoken word production: Evidence from ERPs. *Neuropsychologia, 146*, Article 107559. https://doi.org/10.1016/j.neuropsychologia.2020.107559

Roelofs, A. (2015). Modeling of phonological encoding in spoken word production: From Germanic languages to Mandarin Chinese and Japanese. *Japanese Psychological Research, 57*(1), 22−37.

Schiller, N. O. (2008). The masked onset priming effect in picture naming. *Cognition, 106*(2), 952−962.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming.

Psychological Review, 96(4), 523−568.
Stenneken, P., Conrad, M., & Jacobs, A. M. (2007). Processing of syllables in production and recognition tasks. *Journal of Psycholinguistic Research, 36*(1), 65−78.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276−315.
Tremblay, P., Deschamps, I., Bédard, P., Tessier, M. H., Carrier, M., & Thibeault, M. (2018). Aging of speech production, from articulatory accuracy to motor timing. *Psychology and Aging, 33*(7), 1022−1034.
Varley, R., & Whiteside, S. P. (2001). What is the underlying impairment in acquired apraxia of speech?. *Aphasiology, 15*(1), 39−49.
Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science, 9*(4), 325−329.
Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*(3), 374−408.
Wang, J., Wong, W. K., Wang, S., & Chen, H. C. (2017). Primary phonological planning units in spoken word production are language-specific: Evidence from an ERP study. *Scientific Reports, 7*(1), 5815−5822.
Whiteside, S. P., & Varley, R. A. (1998). A reconceptualisation of apraxia of speech: A synthesis of evidence. *Cortex, 34*(2), 221−231.
You, W. P., Zhang Q. F., & Verdonschot, R. G. (2012). Masked syllable priming effects in word and picture naming in Chinese. *PLoS ONE, 7*(10), Article e46595. https://doi.org/10.1371/journal.pone.0046595
Yu, M. X., Mo, C., & Mo, L. (2014). The role of phoneme in Mandarin Chinese production: Evidence from ERPs. *PLoS ONE, 9*(9), Article e106486. https://doi.org/10.1371/journal.pone.0106486
Zhang, Q. F., & Damian, M. F. (2019). Syllables constitute proximate units for Mandarin speakers: Electrophysiological evidence from a masked priming task. *Psychophysiology, 56*(4), Article e13317. https://doi.org/10.1111/psyp.13317
Zhang, Q. F., & Wang, C. (2014). Syllable frequency and word frequency effects in spoken and written word production in a non-alphabetic script. *Frontiers in Psychology, 5*, Article 120. https://doi.org/10.3389/fpsyg.2014.00120

The cognitive mechanism of syllable frequency effects in speech production: A cross-language perspective

PAN Jiabing, ZHANG Qingfang

(*Department of Psychology, Renmin University of China, Beijing* 100872, *China*)

Abstract: In speech production, the syllable frequency effect refers to the processing advantage of high frequency syllables over low frequency syllables. Features and theories of the syllable frequency effect in speech production of Chinese and Indo-European languages are summarized, and the cross-language differences of the syllable frequency effect are discussed from the perspectives of its main contributors, processing stages, and neural mechanisms. Based on the speech production model and proximate units principle, as well as considering the inherent properties of different languages, we analyse the potential causes for the cross-language differences in syllable frequency effects and propose a model to clarify the mechanism of syllable processing in Chinese speech production, which provides a new perspective on different roles of syllables in speech production of Chinese or alphabetic languages. Future studies should distinguish the effects of token syllable frequency and type syllable frequency, and focus on Chinese syllable frequency effects through a variety of techniques and paradigms, to further explore the cognitive mechanism and cross-language differences of syllable processing in spoken word production.

Keywords: speech production, syllable frequency effect, cross-language, mental syllabary, syllabic neighbours